

PCT

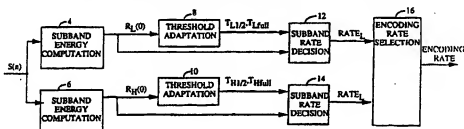
WORLD INTELLECTUAL PROPERTY  
ORGANIZATION

INTERNATIONAL APPLICATION PUBLISHED UNDER



|   |  |  |
|---|--|--|
| (51) International Patent Classification <sup>6</sup> :<br>G10L 3/00, 7/04  |  | (11) Int. No.<br>WO 9605592A1  |
| A1  |  | (43) International Publication Date: 22 February 1996 (22.02.96)   |
| (21) International Application Number: PCT/US95/09830   |  | (81) Designated States: AM, AT, AU, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IS, JP, KE, KG, KP, KR, KZ, LK, LR, LT, LU, LV, MD, MG, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TT, UA, UG, UZ, VN, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG), ARIPO patent (KE, MW, SZ, ZG). |
| (22) International Filing Date: 1 August 1995 (01.08.95)  |  |  |
| (30) Priority Data:<br>288,413 10 August 1994 (10.08.94) US   |  |  |
| (71) Applicant: QUALCOMM INCORPORATED [US/US]; 6455 Lusk Boulevard, San Diego, CA 92121 (US).   |  |  |
| (72) Inventors: DEJACO, Andrew, P.; 10424 Flanders Cove, San Diego, CA 92126 (US). GARDNER, William, R.; 4650 Exbury Court, San Diego, CA 92130 (US). |  |  |
| (74) Agent: MILLER, Russell, B.; Qualcomm Incorporated, 6455 Lusk Boulevard, San Diego, CA 92121 (US).  |  | <p><b>Published</b></p> <p><i>With international search report.</i></p> <p><i>Before the expiration of the time limits for amending the claims and to be republished in the event of the receipt of amendments.</i></p>  |

(54) Title: METHOD AND APPARATUS FOR SELECTING AN ENCODING RATE IN A VARIABLE RATE VOCODER



(57) Abstract

The present invention provides a method by which to reduce the probability of coding low energy unvoiced speech as background noise. An encoding rate is determined by dividing the input signal into subbands using digital subband filters (4) and (6) and comparing the energy in those bands to a set of thresholds in subband rate decision elements (12) and (14) and then examining those comparisons in an encoding rate selector (16). By this method, unvoiced speech can be distinguished from background noise. The present invention, also, provides a means for setting the threshold levels using the signal to noise ratio of the input signal, and the present invention provides a method for coding music through a variable rate vocoder by examining the periodicity of the input signal to distinguish the music from background noise.

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

|    |                          |    |                                       |    |                          |
|----|--------------------------|----|---------------------------------------|----|--------------------------|
| AT | Austria                  | GB | United Kingdom                        | MR | Mauritania               |
| AU | Australia                | GE | Georgia                               | MW | Malawi                   |
| BB | Barbados                 | GN | Guinea                                | NE | Niger                    |
| BE | Belgium                  | GR | Greece                                | NL | Netherlands              |
| BF | Burkina Faso             | HU | Hungary                               | NO | Norway                   |
| BG | Bulgaria                 | IE | Ireland                               | NZ | New Zealand              |
| RJ | Brazil                   | IT | Italy                                 | PL | Poland                   |
| BR | Brazil                   | JP | Japan                                 | PT | Portugal                 |
| BY | Belarus                  | KE | Kenya                                 | RO | Romania                  |
| CA | Canada                   | KG | Kyrgyzstan                            | RU | Russian Federation       |
| CF | Central African Republic | KP | Democratic People's Republic of Korea | SD | Sudan                    |
| CG | Congo                    | KN | Republic of Korea                     | SE | Sweden                   |
| CH | Switzerland              | KZ | Kazakhstan                            | SI | Slovenia                 |
| CI | Côte d'Ivoire            | LJ | Liechtenstein                         | SK | Slovakia                 |
| CM | Cameroon                 | LK | Sri Lanka                             | SN | Senegal                  |
| CN | China                    | LU | Luxembourg                            | TD | Chad                     |
| CS | Czechoslovakia           | LV | Latvia                                | TG | Togo                     |
| CZ | Czech Republic           | MC | Monaco                                | TJ | Tajikistan               |
| DE | Germany                  | MD | Republic of Moldova                   | TT | Trinidad and Tobago      |
| DK | Denmark                  | MG | Madagascar                            | UA | Ukraine                  |
| ES | Spain                    | ML | Mali                                  | US | United States of America |
| FI | Finland                  | MN | Mongolia                              | UZ | Uzbekistan               |
| FR | France                   |    |                                       | VN | Viet Nam                 |
| GA | Gabon                    |    |                                       |    |                          |

## METHOD AND APPARATUS FOR SELECTING AN ENCODING RATE IN A VARIABLE RATE VOCODER

### BACKGROUND OF THE INVENTION

5

#### I. Field of the Invention

The present invention relates to vocoders. More particularly, the present invention relates to a novel and improved method for determining  
10 speech encoding rate in a variable rate vocoder.

#### II. Description of the Related Art

Variable rate speech compression systems typically use some form of  
15 rate determination algorithm before encoding begins. The rate determination algorithm assigns a higher bit rate encoding scheme to segments of the audio signal in which speech is present and a lower rate encoding scheme for silent segments. In this way a lower average bit rate will be achieved while the voice quality of the reconstructed speech will  
20 remain high. Thus to operate efficiently a variable rate speech coder requires a robust rate determination algorithm that can distinguish speech from silence in a variety of background noise environments.

One such variable rate speech compression system or variable rate vocoder is disclosed in copending U.S. Patent Application Serial No.  
25 07/713,661 filed June 11, 1991, entitled "Variable Rate Vocoder" and assigned to the assignee of the present invention, the disclosure of which is incorporated by reference. In this particular implementation of a variable rate vocoder, input speech is encoded using Code Excited Linear Predictive Coding (CELP) techniques at one of several rates as determined by the level  
30 of speech activity. The level of speech activity is determined from the energy in the input audio samples which may contain background noise in addition to voiced speech. In order for the vocoder to provide high quality voice encoding over varying levels of background noise, an adaptively adjusting threshold technique is required to compensate for the affect of  
35 background noise on the rate decision algorithm.

Vocoders are typically used in communication devices such as cellular telephones or personal communication devices to provide digital signal compression of an analog audio signal that is converted to digital form for transmission. In a mobile environment in which a cellular  
40 telephone or personal communication device may be used, high levels of

background noise energy make it difficult for the rate determination algorithm to distinguish low energy unvoiced sounds from background noise silence using a signal energy based rate determination algorithm. Thus unvoiced sounds frequently get encoded at lower bit rates and the voice quality becomes degraded as consonants such as "s","x","ch","sh","t", etc. are lost in the reconstructed speech.

Vocoders that base rate decisions solely on the energy of background noise fail to take into account the signal strength relative to the background noise in setting threshold values. A vocoder that bases its threshold levels solely on background noise tends to compress the threshold levels together when the background noise rises. If the signal level were to remain fixed this is the correct approach to setting the threshold levels, however, were the signal level to rise with the background noise level, then compressing the threshold levels is not an optimal solution. An alternative method for setting threshold levels that takes into account signal strength is needed in variable rate vocoders.

A final problem that remains arises during the playing of music through background noise energy based rate decision vocoders. When people speak, they must pause to breathe which allows the threshold levels to reset to the proper background noise level. However, in transmission of music through a vocoder, such as arises in music-on-hold conditions, no pauses occur and the threshold levels will continue rising until the music starts to be coded at a rate less than full rate. In such a condition the variable rate coder has confused music with background noise.

## SUMMARY OF THE INVENTION

The present invention is a novel and improved method and apparatus for determining an encoding rate in a variable rate vocoder. It is a first objective of the present invention to provide a method by which to reduce the probability of coding low energy unvoiced speech as background noise. In the present invention, the input signal is filtered into a high frequency component and a low frequency component. The filtered components of the input signal are then individually analyzed to detect the presence of speech. Because unvoiced speech has a high frequency component its strength relative to a high frequency band is more distinct from the background noise in that band than it is compared to the background noise over the entire frequency band.

## 3

A second objective of the present invention is to provide a means by which to set the threshold levels that takes into account signal energy as well as background noise energy. In the present invention, the setting of voice detection thresholds is based upon an estimate of the signal to noise ratio (SNR) of the input signal. In the exemplary embodiment, the signal energy is estimated as the maximum signal energy during times of active speech and the background noise energy is estimated as the minimum signal energy during times of silence.

A third objective of the present invention is to provide a method for coding music passing through a variable rate vocoder. In the exemplary embodiment, the rate selection apparatus detects a number of consecutive frames over which the threshold levels have risen and checks for periodicity over that number of frames. If the input signal is periodic this would indicate the presence of music. If the presence of music is detected then the thresholds are set at levels such that the signal is coded at full rate.

### BRIEF DESCRIPTION OF THE DRAWINGS

The features, objects, and advantages of the present invention will become more apparent from the detailed description set forth below when taken in conjunction with the drawings in which like reference characters identify correspondingly throughout and wherein:

Figure 1 is a block diagram of the present invention.

### DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

Referring to Figure 1 the input signal,  $S(n)$ , is provided to subband energy computation element 4 and subband energy computation element 6. The input signal  $S(n)$  is comprised of an audio signal and background noise. The audio signal is typically speech, but it may also be music. In the exemplary embodiment,  $S(n)$  is provided in twenty millisecond frames of 160 samples each. In the exemplary embodiment, input signal  $S(n)$  has frequency components from 0 kHz to 4 kHz, which is approximately the bandwidth of a human speech signal.

In the exemplary embodiment, the 4 kHz input signal,  $S(n)$ , is filtered into two separate subbands. The two separate subbands lie between 0 and 2 kHz and 2 kHz and 4 kHz respectively. In an exemplary embodiment, the input signal may be divided into subbands by subband filters, the design of

which are well known in the art and detailed in U.S. Patent Application Serial No. 08/189,819 filed February 1, 1994, entitled "Frequency Selective Adaptive Filtering", and assigned to the assignee of the present invention, incorporated by reference herein.

- 5 The impulse responses of the subband filters are denoted  $h_L(n)$ , for the lowpass filter, and  $h_H(n)$ , for the highpass filter. The energy of the resulting subband components of the signal can be computed to give the values  $R_L(0)$  and  $R_H(0)$ , simply by summing the squares of the subband filter output samples, as is well known in the art.
- 10 In a preferred embodiment, when input signal  $S(n)$  is provided to subband energy computation element 4, the energy value of the low frequency component of the input frame,  $R_L(0)$ , is computed as:

$$R_L(0) = R_S(0) \cdot R_{hL}(0) + 2 \cdot \sum_{i=1}^{L-1} R_S(i) \cdot R_{hL}(i), \quad (1)$$

- 15 where  $L$  is the number taps in the lowpass filter with impulse response  $h_L(n)$ ,  
 where  $R_S(i)$  is the autocorrelation function of the input signal,  $S(n)$ , given by the equation:

$$R_S(i) = \sum_{n=1}^N S(n) \cdot S(n-i), \quad \text{for } i \in [0, L-1] \quad (2)$$

- where  $N$  is the number of samples in the frame,  
 and where  $R_{hL}$  is the autocorrelation function of the lowpass filter  $h_L(n)$   
 25 given by:

$$R_{hL}(i) = \sum_{n=0}^{L-1} h_L(n) \cdot h_L(n-i), \quad \text{for } i \in [0, L-1] \quad (3)$$

$$= 0 \quad \text{else}$$

- 30 The high frequency energy,  $R_H(0)$ , is computed in a similar fashion in subband energy computation element 6.

The values of the autocorrelation function of the subband filters can be computed ahead of time to reduce the computational load. In addition, some of the computed values of  $R_S(i)$  are used in other computations in the

coding of the input signal,  $S(n)$ , which further reduces the net computational burden of the encoding rate selection method of the present invention. For example, the derivation of LPC filter tap values requires the computation of a set of input signal autocorrelation coefficients.

- 5 The computation of LPC filter tap values is well known in the art and is detailed in the abovementioned U.S. Patent Application 08/004,484. If one were to code the speech with a method requiring a ten tap LPC filter only the values of  $RS(i)$  for  $i$  values from 11 to  $L-1$  need to be computed, in addition to those that are used in the coding of the signal, because  $RS(i)$  for  $i$  values from 0 to 10 are used in computing the LPC filter tap values. In the  
10 exemplary embodiment, the subband filters have 17 taps,  $L=17$ .

- Subband energy computation element 4 provides the computed value of  $RL(0)$  to subband rate decision element 12, and subband energy computation element 6 provides the computed value of  $RH(0)$  to subband  
15 rate decision element 14. Rate decision element 12 compares the value of  $RL(0)$  against two predetermined threshold values  $TL1/2$  and  $TLfull$  and assigns a suggested encoding rate,  $RATE_L$ , in accordance with the comparison. The rate assignment is conducted as follows:

- |    |                               |                             |     |
|----|-------------------------------|-----------------------------|-----|
| 20 | $RATE_L = \text{eighth rate}$ | $RL(0) \leq TL1/2$          | (4) |
|    | $RATE_L = \text{half rate}$   | $TL1/2 < RL(0) \leq TLfull$ | (5) |
|    | $RATE_L = \text{full rate}$   | $RL(0) > TLfull$            | (6) |

- Subband rate decision element 14 operates in a similar fashion and selects a  
25 suggest encoding rate,  $RATE_H$ , in accordance with the high frequency energy value  $RH(0)$  and based upon a different set of threshold values  $TH1/2$  and  $THfull$ . Subband rate decision element 12 provides its suggested encoding rate,  $RATE_L$ , to encoding rate selection element 16, and subband rate decision element 14 provides its suggested encoding rate,  $RATE_H$ , to  
30 encoding rate selection element 16. In the exemplary embodiment, encoding rate selection element 16 selects the higher of the two suggest rates and provides the higher rate as the selected ENCODING RATE.

- Subband energy computation element 4 also provides the low frequency energy value,  $RL(0)$ , to threshold adaptation element 8, where the  
35 threshold values  $TL1/2$  and  $TLfull$  for the next input frame are computed. Similarly, subband energy computation element 6 provides the high frequency energy value,  $RH(0)$ , to threshold adaptation element 10, where the threshold values  $TH1/2$  and  $THfull$  for the next input frame are computed.

- Threshold adaptation element 8 receives the low frequency energy value,  $R_L(0)$ , and determines whether  $S(n)$  contains background noise or audio signal. In an exemplary implementation, the method by which threshold adaptation element 8 determines if an audio signal is present is by examining the normalized autocorrelation function NACF, which is given by the equation:

$$NACF = \max_T \frac{\sum_{n=0}^{N-1} e(n) \cdot e(n-T)}{\frac{1}{2} \left[ \sum_{n=0}^{N-1} e^2(n) + \sum_{n=0}^{N-1} e^2(n-T) \right]}, \quad (7)$$

- 10 where  $e(n)$  is the formant residual signal that results from filtering the input signal,  $S(n)$ , by an LPC filter.

- The design of and filtering of a signal by an LPC filter is well known in the art and is detailed in aforementioned U.S. Patent Application 08/004,484.
- 15 The input signal,  $S(n)$  is filtered by the LPC filter to remove interaction of the formants. NACF is compared against a threshold value to determine if an audio signal is present. If NACF is greater than a predetermined threshold value, it indicates that the input frame has a periodic characteristic indicative of the presence of an audio signal such as speech or
- 20 music. Note that while parts of speech and music are not periodic and will exhibit low values of NACF, background noise typically never displays any periodicity and nearly always exhibits low values of NACF.

- If it is determined that  $S(n)$  contains background noise, the value of NACF is less than a threshold value  $TH1$ , then the value  $R_L(0)$  is used to
- 25 update the value of the current background noise estimate  $BGN_L$ . In the exemplary embodiment,  $TH1$  is 0.35.  $R_L(0)$  is compared against the current value of background noise estimate  $BGN_L$ . If  $R_L(0)$  is less than  $BGN_L$ , then the background noise estimate  $BGN_L$  is set equal to  $R_L(0)$  regardless of the value of NACF.

- 30 The background noise estimate  $BGN_L$  is only increased when NACF is less than threshold value  $TH1$ . If  $R_L(0)$  is greater than  $BGN_L$  and NACF is less than  $TH1$ , then the background noise energy  $BGN_L$  is set  $\alpha_1 \cdot BGN_L$ , where  $\alpha_1$  is a number greater than 1. In the exemplary embodiment,  $\alpha_1$  is equal to 1.03.  $BGN_L$  will continue to increase as long as NACF is less than
- 35 threshold value  $TH1$  and  $R_L(0)$  is greater than the current value of  $BGN_L$ .



until  $BGN_L$  reaches a predetermined maximum value  $BGN_{max}$  at which point the background noise estimate  $BGN_L$  is set to  $BGN_{max}$ .

- If an audio signal is detected, signified by the value of NACF exceeding a second threshold value TH2, then the signal energy estimate,  $S_L$ , is updated. In the exemplary embodiment, TH2 is set to 0.5. The value of  $R_L(0)$  is compared against a current lowpass signal energy estimate,  $S_L$ . If  $R_L(0)$  is greater than the current value of  $S_L$ , then  $S_L$  is set equal to  $R_L(0)$ . If  $R_L(0)$  is less than the current value of  $S_L$ , then  $S_L$  is set equal to  $\alpha_2 \cdot S_L$ , again only if NACF is greater than TH2. In the exemplary embodiment,  $\alpha_2$  is set to 0.96.

Threshold adaptation element 8 then computes a signal to noise ratio estimate in accordance with equation 8 below:

$$SNR_L = 10 \cdot \log \left[ \frac{S_L}{BGN_L} \right] \quad (8)$$

- Threshold adaptation element 8 then determines an index of the quantized signal to noise ratio  $ISNR_L$  in accordance with equation 9-12 below:

$$ISNR_L = \text{nint} \left[ \frac{SNR_L - 20}{5} \right], \quad \text{for } 20 < SNR_L < 55, \quad (9)$$

$$\begin{aligned} &= 0, & \text{for } SNR_L \leq 20, & (10) \\ &= 7, & \text{for } SNR_L \geq 55. \end{aligned}$$

- where nint is a function that rounds the fractional value to the nearest integer.

Threshold adaptation element 8, then selects or computes two scaling factors,  $k_{L1/2}$  and  $k_{Lfull}$ , in accordance with the signal to noise ratio index,

- $ISNR_L$ . An exemplary scaling value lookup table is provided in table 1 below:

8

TABLE 1

| ISNRL | $K_{L1/2}$ | $K_{Lfull}$ |
|-------|------------|-------------|
| 0     | 7.0        | 9.0         |
| 1     | 7.0        | 12.6        |
| 2     | 8.0        | 17.0        |
| 3     | 8.6        | 18.5        |
| 4     | 8.9        | 19.4        |
| 5     | 9.4        | 20.9        |
| 6     | 11.0       | 25.5        |
| 7     | 15.8       | 39.8        |

These two values are used to compute the threshold values for rate selection in accordance with the equations below:

5

$$T_{L1/2} = K_{L1/2} \cdot BGNL, \text{ and} \quad (11)$$

$$T_{Lfull} = K_{Lfull} \cdot BGNL, \quad (12)$$

where  $T_{L1/2}$  is low frequency half rate threshold value and  
 10  $T_{Lfull}$  is the low frequency full rate threshold value.

Threshold adaptation element 8 provides the adapted threshold values  $T_{L1/2}$  and  $T_{Lfull}$  to rate decision element 12. Threshold adaptation element 10 operates in a similar fashion and provides the threshold values  $T_{H1/2}$  and  $T_{Hfull}$  to subband rate decision element 14.

The initial value of the audio signal energy estimate  $S$ , where  $S$  can be  $S_L$  or  $S_H$ , is set as follows. The initial signal energy estimate,  $S_{INIT}$ , is set to -18.0 dBm0, where 3.17 dBm0 denotes the signal strength of a full sine wave, which in the exemplary embodiment is a digital sine wave with an amplitude range from -8031 to 8031.  $S_{INIT}$  is used until it is determined  
 20 that an acoustic signal is present.

The method by which an acoustic signal is initially detected is to compare the NACF value against a threshold, when the NACF exceeds the threshold for a predetermined number consecutive frames, then an acoustic  
 25 signal is determined to be present. In the exemplary embodiment, NACF must exceed the threshold for ten consecutive frames. After this condition is met the signal energy estimate,  $S$ , is set to the maximum signal energy in the preceding ten frames.

The initial value of the background noise estimate  $BGNL$  is initially  
 30 set to  $BGN_{max}$ . As soon as a subband frame energy is received that is less

than  $BGN_{max}$ , the background noise estimate is reset to the value of the received subband energy level, and generation of the background noise  $BGN_L$  estimate proceeds as described earlier.

In a preferred embodiment a hangover condition is actuated when  
 5 following a series of full rate speech frames, a frame of a lower rate is detected. In the exemplary embodiment, when four consecutive speech frames are encoded at full rate followed by a frame where ENCODING RATE is set to a rate less than full rate and the computed signal to noise ratios are less than a predetermined minimum SNR, the  
 10 ENCODING RATE for that frame is set to full rate. In the exemplary embodiment the predetermined minimum SNR is 27.5 dB as defined in equation 8.

In the preferred embodiment, the number of hangover frames is a function of the signal to noise ratio. In the exemplary embodiment, the  
 15 number of hangover frames is determined as follows:

$$\#hangover\ frames = 1 \quad 22.5 < SNR < 27.5, \quad (13)$$

$$\#hangover\ frames = 2 \quad SNR \leq 22.5, \quad (14)$$

$$\#hangover\ frames = 0 \quad SNR \geq 27.5. \quad (15)$$

20

The present invention also provides a method with which to detect the presence of music, which as described before lacks the pauses which allow the background noise measures to reset. The method for detecting the presence of music assumes that music is not present at the start of the call.  
 25 This allows the encoding rate selection apparatus of the present invention to properly estimate and initial background noise energy,  $BGN_{init}$ . Because music unlike background noise has a periodic characteristic, the present invention examines the value of NACF to distinguish music from background noise. The music detection method of the present invention  
 30 computes an average NACF in accordance with the equation below:

$$NACF_{AVE} = \frac{1}{T} \sum_{i=1}^T NACF(i), \quad (16)$$

where NACF is defined in equation 7, and

35 where T is the number of consecutive frames in which the estimated value of the background noise has been increasing from an initial background noise estimate  $BGN_{INIT}$ .

If the background noise BGN has been increasing for the predetermined number of frames T and NACFAVE exceeds a predetermined threshold, then music is detected and the background noise  
5 BGN is reset to BGN<sub>init</sub>. It should be noted that to be effective the value T must be set low enough that the encoding rate doesn't drop below full rate. Therefore the value of T should be set as a function of the acoustic signal and BGN<sub>init</sub>.

The previous description of the preferred embodiments is provided  
10 to enable any person skilled in the art to make or use the present invention. The various modifications to these embodiments will be readily apparent to those skilled in the art, and the generic principles defined herein may be applied to other embodiments without the use of the inventive faculty. Thus, the present invention is not intended to be limited to the  
15 embodiments shown herein but is to be accorded the widest scope consistent with the principles and novel features disclosed herein.

**WE CLAIM:**

## CLAIMS

1. An apparatus for determining an encoding rate for a variable  
2 rate vocoder comprising:  
subband energy computation means for receiving an input signal and  
4 determining a plurality of subband energy values in accordance with a  
predetermined subband energy computation format;  
6 rate determination means for receiving said plurality of subband  
energy values and determining said encoding rate in accordance with said  
8 plurality of subband energy values.

2. The apparatus of Claim 1 wherein said subband energy  
2 computation means determines each of said plurality of subband energy  
values in accordance with the equation:

$$\text{subband energy} = R_S(0) \cdot R_{hbp}(0) + 2 \cdot \sum_{i=1}^{L-1} R_S(i) \cdot R_{hbp}(i),$$

- 6 where L is the number taps in the lowpass filter  $h_L(n)$ ,  
8 where  $R_S(i)$  is the autocorrelation function of the input signal,  $S(n)$ , and  
where  $R_{hL}$  is the autocorrelation function of a bandpass filter  $h_{bp}(n)$ .

3. The apparatus of Claim 1 further comprising threshold  
2 computation means disposed between said subband energy computation  
means and said rate determination means for receiving said subband energy  
4 values and for determining a set of encoding rate threshold values in  
accordance with plurality of subband energy values.

4. The apparatus of Claim 3 wherein said threshold computation  
2 means determines a signal to noise ratio value in accordance with said  
plurality of subband energy values.

5. The apparatus of Claim 4 wherein said threshold computation  
2 means determines a scaling value in accordance with said signal to noise  
ratio value.

6. The apparatus of Claim 5 wherein threshold computation  
2 means determines at least one threshold value by multiplying a background  
noise estimate by said scaling value.

7. The apparatus of Claim 1 wherein said rate determination  
2 compares at least one of said plurality of subband energy values with at least  
one threshold value to determine said encoding rate.

8. The apparatus of Claim 6 wherein said rate determination  
2 means compares at least one of said plurality of subband energy values with  
said at least one threshold value to determine said encoding rate.

9. The apparatus of Claim 1 wherein said rate determination  
2 means determines a plurality of suggested encoding rates wherein each  
suggested encoding rate corresponds to each of said plurality of subband  
4 energy values and wherein said rate determination means determines said  
encoding rate in accordance with said plurality of suggested encoding rates.

10. An apparatus for determining an encoding rate for a variable  
2 rate vocoder comprising:  
signal to noise ratio means for receiving an input signal and  
4 determining a signal to noise ratio value in accordance with said input  
signal;  
6 rate determination means for receiving said signal to noise ratio  
value and determining said encoding rate in accordance with said signal to  
8 noise ratio value.

11. An apparatus for determining an encoding rate for a variable  
2 rate vocoder comprising:  
a subband energy calculator that receives an input signal and  
4 determines a plurality of subband energy values in accordance with a  
predetermined subband energy computation format;  
6 a rate selector that receives said plurality of subband energy values  
and selects said encoding rate in accordance with said plurality of subband  
8 energy values.

12. The apparatus of Claim 11 wherein said subband energy  
2 calculator determines each of said plurality of subband energy values in  
accordance with the equation:

$$\text{subband energy} = R_S(0) \cdot R_{hbp}(0) + 2 \cdot \sum_{i=1}^{L-1} R_S(i) \cdot R_{hbp}(i),$$

6

where  $L$  is the number taps in the lowpass filter  $h_L(n)$ ,

8

where  $R_S(i)$  is the autocorrelation function of the input signal,  $S(n)$ , and  
where  $R_{hL}$  is the autocorrelation function of a bandpass filter  $h_{bp}(n)$ .

2

13. The apparatus of Claim 11 further comprising threshold  
calculator disposed between said subband energy calculator and said rate  
selector that receives said subband energy values and determines a set of  
encoding rate threshold values in accordance with plurality of subband  
energy values.

2

14. The apparatus of Claim 13 wherein said threshold calculator  
determines a signal to noise ratio value in accordance with said plurality of  
subband energy values.

2

15. The apparatus of Claim 14 wherein said threshold calculator  
determines a scaling value in accordance with said signal to noise ratio  
value.

2

16. The apparatus of Claim 15 wherein threshold calculator  
determines at least one threshold value by multiplying a background noise  
estimate by said scaling value.

2

17. The apparatus of Claim 11 wherein said rate selector compares  
at least one of said plurality of subband energy values with at least one  
threshold value to determine said encoding rate.

2

18. The apparatus of Claim 16 wherein said rate selector compares  
at least one of said plurality of subband energy values with said at least one  
threshold value to determine said encoding rate.

4

19. The apparatus of Claim 11 wherein said rate selector  
determines a plurality of suggested encoding rates wherein each suggested  
encoding rate corresponds to each of said plurality of subband energy values  
and wherein said rate selector determines said encoding rate in accordance  
with said plurality of suggested encoding rates.

2

20. An apparatus for determining an encoding rate for a variable  
rate vocoder comprising:

- a signal to noise ratio calculator that receives an input signal and  
 4 determining a signal to noise ratio value in accordance with said input  
 signal;  
 6 rate selector that receives said signal to noise ratio value and selects  
 said encoding rate in accordance with said signal to noise ratio value.

21. A method for determining an encoding rate for a variable rate  
 2 vocoder comprising the steps of:  
 receiving an input signal;  
 4 determining a plurality of subband energy values in accordance with  
 a predetermined subband energy computation format; and  
 6 determining said encoding rate in accordance with said plurality of  
 subband energy values.

22. The method of Claim 21 wherein said step of determining a  
 2 plurality of subband energy values is performed in accordance with the  
 equation:

$$4 \quad \text{subband energy} = R_S(0) \cdot R_{h_{bp}}(0) + 2 \cdot \sum_{i=1}^{L-1} R_S(i) \cdot R_{h_{bp}}(i),$$

- 6 where L is the number taps in the lowpass filter  $h_L(n)$ ,  
 8 where  $R_S(i)$  is the autocorrelation function of the input signal,  $S(n)$ , and  
 where  $R_{h_L}$  is the autocorrelation function of a bandpass filter  $h_{bp}(n)$ .

23. The method of Claim 21 further comprising the step of  
 2 determining a set of encoding rate threshold values in accordance with  
 plurality of subband energy values.

24. The method of Claim 23 wherein said step of determining a set  
 2 of encoding rate threshold values determines a signal to noise ratio value in  
 accordance with said plurality of subband energy values.

25. The method of Claim 24 wherein said step of determining a set  
 2 of encoding rate threshold values determines a scaling value in accordance  
 with said signal to noise ratio value.



25. The method of Claim 25 wherein said step of determining a set  
2 of encoding rate threshold values determines said rate threshold value by  
multiplying a background noise estimate by said scaling value.

27. The method of Claim 21 wherein said determining said  
2 encoding rate compares at least one of said plurality of subband energy  
values with at least one threshold value to determine said encoding rate.

28. The method of Claim 26 wherein said step of said determining  
2 said encoding rate compares at least one of said plurality of subband energy  
values with said at least one threshold value to determine said encoding  
4 rate.

29. The method of Claim 21 further comprising the step of  
2 generating a suggested encoding rate in accordance with each of said  
plurality of subband energy values and wherein said step of determining an  
4 encoding rate selects one of said suggested encoding rates.

30. A method for determining an encoding rate for a variable rate  
2 vocoder comprising the steps of:  
receiving an input signal;  
4 determining a signal to noise ratio value in accordance with said  
input signal; and  
6 determining said encoding rate in accordance with said signal to noise  
ratio value.

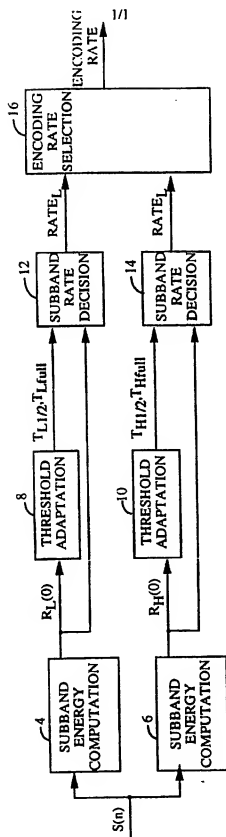


FIG. 1

# INTERNATIONAL SEARCH REPORT

International Application No.  
PCT/US 95/09830

## A. CLASSIFICATION OF SUBJECT MATTER

G 10 L 3/00, G 10 L 7/04

According to International Patent Classification (IPC) or to both national classification and IPC <sup>6</sup>

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G 10 L, H 03 H

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

| Category* | Citation of document, with indication, where appropriate, of the relevant passages   | Relevant to claim No.                      |
|-----------|--|--|
| X         | EP, A, 0 167 364<br>(AMERICAN TELEPHONE AND<br>TELEGRAPH COMPANY)<br>08 January 1986 (08.01.86),<br>fig. 1; abstract;<br>claims 1-9.                           | 1,3-8,<br>10,11,<br>20,21,<br>23-28,<br>30 |
| A         | EP, A, 0 190 796 --<br>(TELECOMMUNICATIONS RADIO-<br>ELECTRIQUES ET TELEPHONIQUES<br>T.R.T.) 13 August 1986<br>(13.08.86),<br>fig. 4; abstract;<br>claims 1-5. | 1-30                                       |
| A         | US, A, 5 054 075 --<br>(HONG et al.) 01 October 1991<br>(01.10.91),<br>fig. 1,2; abstract;   | 1-30                                       |

☒ Further documents are listed in the continuation of box C.

☐ Patent family members are listed in annex.

### \* Special categories of cited documents :

"A" document defining the general state of the art which is not

considered to be of particular relevance

"E" earlier document but published on or after the international

filing date

"L" document which may throw doubts on priority claim(s) or

which is cited to establish the publication date of another

claim or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or

other means

"P" document published prior to the international filing date but

later than the priority date claimed

"T" later document published after the international filing date

or priority date and not in conflict with the application but

used to understand the principle or theory underlying the

invention

"X" document of particular relevance; the claimed invention

cannot be considered novel or cannot be considered to

involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention

cannot be considered to involve an inventive step when the

document is combined with one or more other such docu-

ments, such combination being obvious to a person skilled

in the art.

"A" document member of the same patent family

Date of the actual completion of the international search

10 November 1995

Date of mailing of the international search report

22.12.95

Name and mailing address of the ISA

European Patent Office, P.O. Box 1818 Patenlaan 2

NL - 2280 LV Rijswijk

Tel. (+ 31-70) 340-2040, Tx. 31 651 epo nl,

fax (+ 31-70) 340-3016

Authorized officer

BERGER e.h.

# INTERNATIONAL SEARCH REPORT

International Application No.  
PCT/US 95/09830

| C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT |  |                       |
|--|--|-----------------------|
| Category   | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|  | claims 1-15. -----   |                       |

# ANHANG

zum internationalen Recherchen-  
bericht über die internationale  
Patentanmeldung Nr.

# ANNEX

to the International Search  
Report to the International Patent  
Application No.

# ANNEXE

au rapport de recherche inter-  
national relatif à la demande de brevet  
international n°

PCT/US 95/09830 SAE 115833

In diesem Anhang sind die Mitglieder  
der Patentfamilien der in obenge-  
nannten internationalen Recherchenbericht  
angeführten Patentdokumente angegeben.  
Diese Angaben dienen nur zur Unter-  
richtung und erfolgen ohne Gewähr.

This Annex lists the patent family  
members relating to the patent documents  
cited in the above-mentioned inter-  
national search report. The Office is  
in no way liable for these particulars  
which are given merely for the purpose  
of information.

La présente annexe indique les  
membres de la famille de brevets  
relatifs aux documents de brevets cités  
dans le rapport de recherche inter-  
national visé ci-dessus. Les renseigne-  
ments fournis sont donnés à titre indica-  
tif et n'engagent pas la responsabilité  
de l'Office.

| In Recherchenbericht<br>angeführtes Patentdokument<br>Patent document cited<br>In search report<br>Document de brevet cité<br>dans le rapport de recherche | Datum der<br>Veröffentlichung<br>Publication<br>date<br>Date de<br>publication | Mitglied(er) der<br>Patentfamilie<br>Patent family<br>member(s)<br>Membre(s) de la<br>famille de brevets | Datum der<br>Veröffentlichung<br>Publication<br>date<br>Date de<br>publication |
|--|--|--|--|
| EP A1 167364   | 08-01-86   | JP A2 61032900   | 15-02-86   |
| EP A1 190796   | 13-08-86   | AU A1 52847/86   | 07-08-86   |
|  |  | CA A1 124862   | 10-01-86   |
|  |  | DE CO 3667742  | 25-01-90   |
|  |  | DE EP 190796   | 20-12-89   |
|  |  | FR B1 2577084  | 08-08-86   |
|  |  | FR B1 2577084  | 20-03-86   |
|  |  | JP A2 61177019   | 08-08-86   |
|  |  | US A 4799179   | 17-01-89   |
| US A 5054075   | 01-10-91   | keine - none - rien  |  |